

Performance vs complexity trade-offs for Markovian networked jump estimators

D. Dolz, D. E. Quevedo, I. Peñarrocha, and R. Sanchis

Abstract—This paper addresses the design of a state observer for networked systems with random delays and dropouts. The model of plant and network covers the cases of multiple sensors, out-of-sequence and buffered measurements. The measurement outcomes over a finite interval model the network measurement reception scenarios, which follow a Markov distribution. We present a tractable optimization problem to precalculate off-line a finite set of gains of jump observers. The proposed procedure allows us to trade the complexity of the observer implementation for achieved performance. Several examples illustrate that the on-line computational cost of the observer implementation is lower than that of the Kalman filter, whilst the performance is similar.

I. INTRODUCTION

Networked control systems are control systems where the information (output measurements and/or control inputs) is transmitted via a shared network. The use of networks reduces the installation cost and increases the flexibility, but leads to several network-induced effects such as time delays and packet dropouts (see [5] and [2]). Control and estimation through a network must overcome these problems.

Considering the estimation problem, Kalman filter based solutions may give optimal performance, but at the expense of significant on-line computational complexity. The observer gain is time varying and must be computed online, even for linear time invariant systems (e.g. [6], [12] and [11]). This motivates the search for computationally low cost alternatives. In particular, the use of precalculated gains decreases the need of the implementation computing capacity, but increases the estimation error and requires both storage and a mechanism to choose the appropriate gain at each instant (e.g. [13], [10], [8] and [4]). The jump linear estimator approach proposed in [13] improves the estimation with a set of precalculated gains which are chosen depending on the history of measurement availabilities. A better performance is achieved at the cost of increasing the estimator complexity in terms of storage requirements and gain selection mechanism. An approach of intermediate complexity is presented in recent work [4] where the authors propose a gain dependency on the possible instant and arrival delay for each measurement in a finite set. Computing the gains off-line takes advantage from prior statistical knowledge about the network behavior. When the network behaves as a Markov chain, the design uses the transition probabilities ([13] and [4]).

D. Dolz, I. Peñarrocha, and R. Sanchis are with Department of Industrial System Engineering and Design, Universitat Jaume I of Castelló, Spain {ddolz, ipenarro, rsanchis}@uji.es

D. E. Quevedo is with the School of Electrical Engineering and Computer Science, The University of Newcastle, NSW, Australia dquevedo@ieee.org

In this paper we face the estimator design problem for multisensor systems and networks with induced unbounded time-varying delays with known distribution. We derive a finite measurement outcomes parameter that models the network effects and follows a finite Markov chain. Based on this process, we propose a jump linear estimator that gives favorable trade-offs between on-line computational burden and estimation performance. Furthermore, we analyze the effects of reducing the number of stored gains (i.e., complexity) by means of sharing the use of each gain for different values of the finite measurement outcomes parameter.

Two are the main contributions of our current work with respect to [13] and [4]. First, we consider the multisensor with multiple delays scenario. Second, we introduce a flexible way to handle different strategies for the gain dependency to find a compromise between implementation cost and estimation performance. Moreover, the measurement reception model derived here allows to handle more complex gain observer dependencies that cannot be included in [4]. The present work differs from our recent manuscript [8] mainly in the consideration of the stochastic network behavior with unbounded consecutive dropouts instead of a deterministic approach.

The paper has the following structure. In Section II we describe the process, model the network effects, present the observer algorithm and derive estimation error expressions. In Section III we develop the observer design, and demonstrate its convergence. In Section IV we show how gain grouping approaches can be used to find a compromise between implementation cost and performance. Simulation studies are given in Section V, and Section VI draws conclusions.

II. PROBLEM APPROACH

Let us consider linear time invariant discrete-time systems of the form

$$x[t+1] = Ax[t] + B_u u[t] + B_w w[t], \quad (1)$$

$$y_s[t] = c_s x[t] + v_s[t], \quad (2)$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^{n_u}$ is the control input, $y_s \in \mathbb{R}$ is the s -th measured output ($s = 1, \dots, n_y$) with $y[t] = [y_1[t] \dots y_{n_y}[t]]^T$, $w \in \mathbb{R}^{n_w}$ is the state disturbance modeled as a white noise signal of zero mean and known covariance $\mathbf{E}\{w[t]w[t]^T\} = W$, and $v_s \in \mathbb{R}$ is the s -th sensor noise assumed as an independent zero mean white noise signal with known variance $\mathbf{E}\{v_s[t]^2\} = \sigma_s^2$. Throughout this work we assume that the control input is causally available at all times, see Fig. 1.

Let us assume that samples from several sensors are taken synchronously with the input update and sent independently

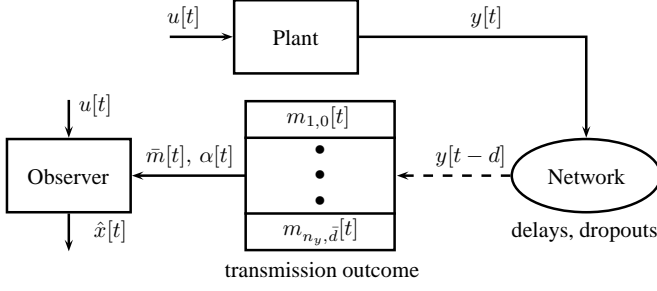


Fig. 1. Networked state estimator.

to the estimator unit through a network with packet dropouts and induced time-varying delays (see Fig. 1). Let us assume synchronization between sensors and the estimator unit and time-tagged message sending. We denote as $\tau_s[t] \in \mathbb{N}$ the induced delay on the delivery of the t -th sample of sensor s , where $\tau_s[t] = \infty$ represents a measurement loss. We assume that the delays are bounded by \bar{d} ; otherwise, we discard the measurement. Then, the network induced delay for all sensors can take values in a finite range $\tau[t] \in \{0, 1, \dots, \bar{d}\} \cup \infty$. The available information at instant t at the estimator unit is the pair $(m_{s,d}[t], \alpha_{s,d}[t])$ for all $s = 1, \dots, n_y$, $d = 0, 1, \dots, \bar{d}$ being the induced delay, where

$$m_{s,d}[t] = \alpha_{s,d}[t] y_s[t-d], \quad (3)$$

and

$$\alpha_{s,d}[t] = \begin{cases} 1 & \text{if } y_s[t-d] \text{ is received at instant } t \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

Note that $\alpha_{s,d}[t] = 1$ represents that the induced delay of the measurement from sensor s sampled at instant $t-d$ is $\tau_s[t-d] = d$. We consider $m_{s,d}[t] = 0$ if $y_s[t-d]$ does not arrive at instant t .

We introduce an aggregated model to deal with delayed measurements. The aggregated model including the delayed states is

$$\bar{x}[t+1] = \bar{A}\bar{x}[t] + \bar{B}_u u[t] + \bar{B}_w w[t], \quad (5)$$

where $\bar{x}[t] = [x[t]^T \dots x[t-\bar{d}]^T]^T$ and

$$\bar{A} = \begin{bmatrix} A & 0 & \dots & 0 \\ I & \dots & 0 & \vdots \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & I & 0 \end{bmatrix}, \quad \bar{B}_u = \begin{bmatrix} B_u \\ 0 \end{bmatrix}, \quad \bar{B}_w = \begin{bmatrix} B_w \\ 0 \end{bmatrix},$$

We incorporate \bar{d} additional fictitious sensors for each actual sensor with a different constant delay, and express the available measurements from real sensor s as

$$\bar{m}_s[t] = [m_{s,0}[t] \dots m_{s,\bar{d}}[t]]^T, \quad (6)$$

with $m_{s,d}[t]$ as defined in (3). With that, the number of total (real and fictitious) sensors is $\bar{n}_y = n_y(1 + \bar{d})$. This model handles out-of-sequence and buffered samples (see [9]).

A. Network modeling

Let us define the process $\theta[t]$ which captures the measurement transmission outcomes at times $\{t-\bar{d}, \dots, t\}$ as follows:

$$\theta[t] = [\theta_1[t] \dots \theta_{n_y}[t]]^T, \quad (7)$$

with $\theta[t]$ a binary column vector of length $n_\theta = \frac{(\bar{d}+1)(\bar{d}+2)}{2} n_y$ and where $\theta_s[t]$ represents the measurement reception at times $\{t-\bar{d}, \dots, t\}$ from sensor s (cf. [4])

$$\theta_s[t] = [\theta_{s,0}[t] \dots \theta_{s,\bar{d}}[t]], \quad (8)$$

$$\theta_{s,d}[t] = [\alpha_{s,0}[t-d] \quad \alpha_{s,1}[t-d+1] \quad \dots \quad \alpha_{s,d}[t]]. \quad (9)$$

$\theta_{s,d}[t]$ represents the transmission outcome of measurement $y_s[t-d]$ at times $\{t-\bar{d}, \dots, t\}$. In real-time systems, measurement $y_s[t-d]$ can only be received once. This implies that $\|\theta_{s,d}[t]\|_1 \leq 1$ with $\|\theta_{s,d}[t]\|_1 = \sum_{d=0}^{\bar{d}} \alpha_{s,d}[t-d+d]$. Clearly, $\theta[t]$ is an ergodic¹ Markov chain (see [1]) that can take values in the finite set

$$\theta[t] \in \Theta = \{\vartheta_0, \vartheta_1, \dots, \vartheta_r\}, \quad r = ((\bar{d}+2)!)^{n_y} - 1, \quad (10)$$

and where ϑ_i (for $i = 0, \dots, r$) denotes each possible combination of the historical measurement transmission outcomes. $\vartheta_0 = 0$ denotes the case where neither of the samples from $t-\bar{d}$ to t is received. To obtain the transition probabilities, we use the following assumption.

Assumption 1: The delays $\tau_s[t]$ are i.i.d random variables with $\beta_{s,d} \triangleq \Pr\{\tau_s[t] = d\}^2$ for $d = 0, \dots, \bar{d}$, and where $\sum_{d=0}^{\bar{d}} \beta_{s,d} \leq 1$. \square

Let us denote the tail probabilities as $\bar{\beta}_{s,d} = \Pr\{\tau_s[t] > d\} = \Pr\{\|\theta_{s,d}[t]\|_1 = 0\}$. Using Assumption 1, the elements from the transition probability matrix $\Lambda = [p_{i,j}]$ with $p_{i,j} = \Pr\{\theta[t+1] = \vartheta_j | \theta[t] = \vartheta_i\}$ are calculated as

$$\prod_{s=1}^{n_y} g(t+1, \bar{d}, s) / g(t, \bar{d}-1, s), \quad (11)$$

where

$$g(t, \bar{d}, s) = \prod_{\substack{d=0 \\ d: \|\theta_{s,d}[t+1]\|_1=0}}^{\bar{d}} \bar{\beta}_{s,d} \prod_{\substack{d=0 \\ d: \alpha_{s,d}[t+1]=1}}^{\bar{d}} \beta_{s,d}. \quad (12)$$

Equation (11) is only valid for feasible transitions. A transition is feasible if $\alpha_{s,d}[t-h]$ have the same value in both $\theta[t]$ and $\theta[t+1]$, for all $s = 1, \dots, n_y$, $d = 0, \dots, \bar{d}-1$ and $h = 0, \dots, \bar{d}-1-d$.

Moreover, we denote the total probability of being at a given state i as $\pi_i[t] = \Pr\{\theta[t] = \vartheta_i\}$, where $\pi[t] = [\pi_1[t], \dots, \pi_r[t]]$ and $\pi[t+1] = \pi[t]\Lambda$.

Let us now define the measurement availability matrix at instant t as

$$\alpha[t] = f(\theta[t]) = \bigoplus_{s=1}^{n_y} \left(\bigoplus_{d=0}^{\bar{d}} \alpha_{s,d}[t] \right), \quad (13)$$

¹In an ergodic Markov chains every state can be reached from every state in a finite time.

²Note that $\Pr\{\alpha_{s,d}[t] = 1\} = \Pr\{\tau_s[t] = d\} = \beta_{s,d}$.

where \oplus denotes the direct sum³. The possible values of $\alpha[t]$ are within a known set

$$\alpha[t] \in \Xi = \{\eta_0, \eta_1, \dots, \eta_q\}, \quad (14)$$

where η_i (for $i = 1, \dots, q$) denotes each possible combination, being η_0 the scenario without available measurements, (i.e., $\eta_0 = 0$). In the general case, any combination of available sensor measurement and delay is possible, leading to $q = 2^{\bar{n}_y} - 1$. $\alpha[t]$ is the result of applying a surjective function $f : \Theta \rightarrow \Xi$ on $\theta[t]$ meaning that $\alpha[t]$ is not a Markov variable neither i.i.d.

Example 1: Let us consider a system with one sensor and $\bar{d} = 1$. Then $\Theta = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$. Fig. 2 illustrates the relationship between θ_t and θ_{t+1} . $\theta_t = \vartheta_0$ means that $y[t]$ has not arrived at t (but can still arrive, i.e. $\tau[t] > 0$) and that $y[t-1]$ is lost ($\tau[t-1] > 1$). $\Pr\{\theta[t+1] = \vartheta_2 | \theta[t] = \vartheta_0\} = 0$ because $\theta[t+1] = \vartheta_2$ would imply that $\tau[t] = 0$, and $\theta[t] = \vartheta_0$ guarantees that $\tau[t] > 0$. However,

$$\begin{aligned} \Pr\{\theta[t+1] = \vartheta_1 | \theta[t] = \vartheta_0\} \\ &= \Pr\{\tau[t+1] = 0, \tau[t] > 1 | \tau[t] > 0, \tau[t-1] > 1\} \\ &= \Pr\{\tau[t+1] = 0, \tau[t] > 1 | \tau[t] > 0\} \\ &= \Pr\{\tau[t+1] = 0\} \Pr\{\tau[t] > 1 | \tau[t] > 0\} \\ &= \Pr\{\alpha_0[t+1] = 1\} \Pr\{\|\theta_1[t+1]\|_1 = 0 | \|\theta_0[t]\|_1 = 0\} \\ &= \Pr\{\alpha_0[t+1] = 1\} \Pr\{\|\theta_1[t+1]\|_1 = 0\} / \Pr\{\|\theta_0[t]\|_1 = 0\} \\ &= \beta_0 \bar{\beta}_1 / \bar{\beta}_0. \end{aligned}$$

The full transition matrix can be obtained in the same way, leading to

$$\begin{bmatrix} \bar{\beta}_1 & \beta_0 \bar{\beta}_1 / \bar{\beta}_0 & 0 & 0 & \beta_1 & \beta_0 \beta_1 / \bar{\beta}_0 \\ 0 & 0 & \bar{\beta}_0 & \beta_0 & 0 & 0 \\ \bar{\beta}_1 & \beta_0 \bar{\beta}_1 / \bar{\beta}_0 & 0 & 0 & \beta_1 & \beta_0 \beta_1 / \bar{\beta}_0 \\ 0 & 0 & \bar{\beta}_0 & \beta_0 & 0 & 0 \\ \bar{\beta}_1 & \beta_0 \bar{\beta}_1 / \bar{\beta}_0 & 0 & 0 & \beta_1 & \beta_0 \beta_1 / \bar{\beta}_0 \\ 0 & 0 & \bar{\beta}_0 & \beta_0 & 0 & 0 \end{bmatrix}.$$

In this case,

$$\Xi = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\} = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\},$$

where only the diagonal terms of η_i have been represented. \square

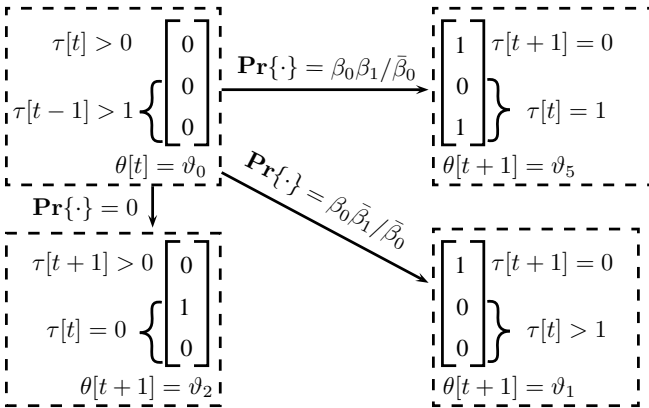


Fig. 2. Markovian transitions example.

Using $\alpha[t]$, we rewrite the received measurement information at instant t as

$$\bar{m}[t] = \alpha[t] (\bar{C}\bar{x}[t] + \bar{v}[t]) \quad (15)$$

³The direct sum between of two matrices, i.e. $A \oplus B$, creates a block diagonal matrix with A and B on the diagonal.

with $\bar{m}[t] = [\bar{m}_1[t]^T \dots \bar{m}_{n_y}[t]^T]^T$, $\bar{v}_t = [\bar{v}_1[t] \dots \bar{v}_{n_y}[t]]^T$ and $\bar{v}_1[t] = [v_1[t] \dots v_1[t]]^T$. The rows of \bar{C} are $\bar{c}_s = [\bar{c}_{s,0} \dots \bar{c}_{s,\bar{d}}]^T$ with $\bar{c}_{s,d} = [0_{1 \times n \cdot d} \ c_s \ 0_{1 \times n \cdot (\bar{d}-d)}]^T$. In (15), $\bar{v}[t]$ is the measurement noise vector with covariance $\mathbf{E}\{\bar{v}[t]\bar{v}[t]^T\} = V = \bigoplus_{s=1}^{n_y} \left(\bigoplus_{d=0}^{\bar{d}} \sigma_s^2 \right)$.

B. Proposed observer

Let us represent $x[t]$ as \hat{x}_t . We propose the following state estimation algorithm. At each instant t , the model is run in open loop leading to the prior estimation

$$\hat{\hat{x}}_{t-} = \bar{A} \hat{\hat{x}}_{t-1} + \bar{B}_u u_{t-1}. \quad (16)$$

If no measurement is received, the best estimation of the system state is the prior estimation, i.e., $\hat{\hat{x}}_t = \hat{\hat{x}}_{t-}$. Otherwise, the estimation state is updated as

$$\hat{\hat{x}}_t = \hat{\hat{x}}_{t-} + L[t](\bar{m}_t - \alpha_t \bar{C} \hat{\hat{x}}_{t-}), \quad (17)$$

where $L[t]$ is the updating gain matrix.

Considering (5) and (15)-(17), the dynamic of the estimation error, defined as $\tilde{x}_t = \hat{x}_t - \hat{\hat{x}}_t$, is

$$\tilde{x}_t = (I - L[t]\alpha_t \bar{C}) (\bar{A} \tilde{x}_{t-1} + \bar{B}_w w_{t-1}) - L[t]\alpha_t \bar{v}_t. \quad (18)$$

The aim of this work is to compute the gain matrices $L[t]$ that minimize the state estimation error while requiring low computing and storage capabilities. Thus, we propose to relate the gains with θ_t as $L[t] = L(\theta_t)$.

In the motivating example in [13], the authors showed that the gains obtained with a Kalman filter depend on the history of combination of sensor availability. In the present work we extend their result to delayed measurements and multisensor transmission defining the gains as

$$L(\theta_t) = \begin{cases} 0 & \text{if } f(\theta_t) = \eta_0, \text{ (no measurement received)} \\ L_i & \text{if } \theta_t = \vartheta_i, f(\vartheta_i) \neq \eta_0 \end{cases} \quad (19)$$

The matrices are computed off-line leading to the finite set

$$L(\theta_t) \in \mathcal{L} = \{L_0, \dots, L_r\}. \quad (20)$$

We will next show how to design such an observer when imposing constraints over \mathcal{L} .

III. OBSERVER DESIGN

As the Markov chain $\{\theta_t\}$ is ergodic, it has a stationary distribution which satisfies $\pi = \pi\Lambda$. We assume the initial condition $\pi[0] = \pi$, and in consequence $\pi[t] = \pi \forall t$. Based on this assumption, the following theorem expresses the evolution of the state estimation error covariance matrix.

Theorem 1: Let $P_{t-1,i} = \mathbf{E}\{\tilde{x}_{t-1}\tilde{x}_{t-1}^T | \theta[t-1] = \vartheta_i\}$ (with $i = 1, \dots, r$) be the covariance matrix for the state estimation error updated at the measurement instant $t-1$ with information $\theta[t-1] = \vartheta_i$. The expected value of the covariance matrix at the measurement instant t , $\mathbf{E}\{\tilde{x}_t \tilde{x}_t^T\}$, is given by

$$\sum_{j=0}^r \mathbf{E}\{\tilde{x}_t \tilde{x}_t^T | \theta[t] = \vartheta_j\} \Pr\{\theta[t] = \vartheta_j\} = \sum_{j=0}^r P_{t,j} \pi_j, \quad (21)$$

where $P_{t,j}$ is defined by

$$\sum_{i=0}^r p_{i,j} \frac{\pi_i}{\pi_j} (F_j (\bar{A} P_{t-1,i} \bar{A}^T + \bar{B}_w W \bar{B}_w^T) F_j^T + X_j V X_j^T), \quad (22)$$

with

$$F_j = I - L_j f(\theta_j) \bar{C}, \quad X_j = L_j f(\theta_j). \quad (23)$$

Proof: See Appendix A ■

The previous theorem establishes a recursion for the covariance matrix. We thus write $\mathcal{P}_t = \mathfrak{E}\{\mathcal{P}_{t-1}\}$, where $\mathcal{P}_t \triangleq (P_{t,0}, \dots, P_{t,r})$, $\mathfrak{E}\{\cdot\} \triangleq (\mathfrak{E}_0\{\cdot\}, \dots, \mathfrak{E}_r\{\cdot\})$, being $\mathfrak{E}_i\{\cdot\}$ the linear operator that returns equation (22). In order to compute the observer gains off-line, one must find the stable solution to the Riccati equation $\mathfrak{E}\{\mathcal{P}_{t-1}\} = \mathcal{P}_{t-1}$. In general, for cases where the observer gain depends on each state of the Markov chain, an explicit expression on the observer gain values can be found using the methods of [13] and [4]. However, the methods applied in those works become untractable for the design of an observer that share the same gain for different states of the Markov chain. Hence, those methods do not directly allow to explore trade-offs between storage complexity and estimation performance. To address this issue, we adopt the following alternative optimization problem

$$\min_{\mathcal{L}, \mathcal{P}} \text{tr} \left(\sum_{j=0}^r P_j \pi_j \right) \quad (24a)$$

$$\text{s.t.} \quad \mathfrak{E}\{\mathcal{P}\} - \mathcal{P} \preceq 0, \quad (24b)$$

with $\mathcal{P} \triangleq (P_0, \dots, P_r)$.

As we shall see next, the constraint in (24b) is instrumental for guaranteeing boundedness of $\mathbf{E}\{\tilde{x}_t \tilde{x}_t^T\}$, and therefore stochastic stability. Note that the next results are independent on the constraints over \mathcal{L} .

A. Boundedness of the covariance

We show in the following that if we apply the gains \mathcal{L} obtained from problem (24), then the sequence $\{\mathcal{P}_t\}$ (and thus $\{\mathbf{E}\{x_t x_t^T\}\}$) converges to the unique solution $\bar{\mathcal{P}} \triangleq (\bar{P}_1, \dots, \bar{P}_r)$ obtained in (24).

Let us first introduce the following lemma, extended from [12], where $\bar{\mathcal{P}} \succ 0$ denotes $\bar{P}_i \succ 0, \forall i = 1 \dots, r$.

Lemma 1: Define the linear operator

$$\mathcal{T}_j(\mathcal{Y}) = \sum_{i=0}^r p_{i,j} \frac{\pi_i}{\pi_j} F_j \bar{A} Y_i \bar{A}^T F_j^T$$

where $\mathcal{T}(\cdot) \triangleq (\mathcal{T}_0(\cdot), \dots, \mathcal{T}_r(\cdot))$ and $\mathcal{Y} \triangleq (Y_0, \dots, Y_r)$. Suppose that there exists $\bar{\mathcal{Y}} \triangleq (\bar{Y}_0, \dots, \bar{Y}_r) \succ 0$ such that $\mathcal{T}(\bar{\mathcal{Y}}) \prec \bar{\mathcal{Y}}$. Then, (a) for all $\mathcal{W} \triangleq (W_0, \dots, W_r) \succeq 0$, $\lim_{t \rightarrow \infty} \mathcal{T}^t(\mathcal{W}) = 0^4$; (b) let $U \succeq 0$ and consider the linear system $\mathcal{Y}_{t+1} = \mathcal{T}(\mathcal{Y}_t) + U$, initialized at \mathcal{Y}_0 , then the sequence $\{\mathcal{Y}_t\}$ is bounded. ■

Using the above lemma, the following theorem proves the boundedness of $\{\mathcal{P}_t\}$.

⁴ $\mathcal{T}^t\{\cdot\}$ represents the recursion of $\mathcal{T}\{\cdot\}$.

Theorem 2: Under Assumption 1, suppose that the set \mathcal{L} in (20) fulfills restriction (24b), i.e., there exists $\bar{\mathcal{P}} \succ 0$ such that $\mathfrak{E}\{\bar{\mathcal{P}}\} \preceq \bar{\mathcal{P}}$. Then, for any initial condition $\mathcal{P}_0 \succeq 0$ the sequence $\{\mathcal{P}_t\}$ is bounded, i.e., $\{\mathcal{P}_t\} \preceq M_{\mathcal{P}}$, with $M_{\mathcal{P}} \triangleq (M_{P_0}, \dots, M_{P_r})$.

Proof: See Appendix A ■

By means of the previous theorem, the next result establishes that $\{\mathcal{P}_t\}$ converges to the solution of problem (24).

Theorem 3: Under Assumption 1, suppose that the set \mathcal{L} in (20) solves problem (24). Then, for any initial condition $\mathcal{P}_0 \succeq 0$, the iteration $\mathcal{P}_{t+1} = \mathfrak{E}\{\mathcal{P}_t\}$ converges to the unique positive semi-definite solution $\bar{\mathcal{P}}$ obtained in problem (24), i.e., $\lim_{t \rightarrow \infty} \mathcal{P}_t = \lim_{t \rightarrow \infty} \mathfrak{E}^t\{\mathcal{P}_0\} = \bar{\mathcal{P}} \succeq 0$, where $\bar{\mathcal{P}} = \mathfrak{E}\{\bar{\mathcal{P}}\}$.

Proof: See Appendix A ■

B. Numerical issues

Problem (24) can be solved using the following linear matrix inequalities and bilinear equality constraints,

$$\min_{\mathcal{L}, \mathcal{P}, \mathcal{R}} \text{tr} \left(\sum_{j=0}^r P_j \pi_j \right) \quad (25a)$$

$$\begin{bmatrix} P_j & \bar{M}_j \bar{A} & \bar{M}_j \bar{W} & \bar{X}_j \bar{V} \\ \bar{A}^T \bar{M}_j^T & \bar{R} & 0 & 0 \\ \bar{W}^T \bar{M}_j^T & 0 & \bar{W} & 0 \\ \bar{V}^T \bar{X}_j^T & 0 & 0 & \bar{V} \end{bmatrix} \succeq 0, \forall j = 0, \dots, r \quad (25b)$$

$$\bar{P} \bar{R} = I \quad (25c)$$

with

$$\begin{aligned} \bar{X}_j &= \left[\sqrt{p_{0,j} \pi_0 / \pi_j} L_j f(\vartheta_j) \cdots \sqrt{p_{r,j} \pi_r / \pi_j} L_j f(\vartheta_j) \right], \\ \bar{M}_j &= \left[\sqrt{p_{0,j} \pi_0 / \pi_j} F_j \cdots \sqrt{p_{r,j} \pi_r / \pi_j} F_j \right], \quad \bar{A} = \bigoplus_{i=0}^r \bar{A}, \\ \bar{W} &= \bigoplus_{j=0}^r \bar{B}_w W \bar{B}_w^T, \quad \bar{V} = \bigoplus_{j=0}^r V, \quad \bar{R} = \bigoplus_{j=0}^r R_j, \quad \bar{P} = \bigoplus_{j=0}^r P_j, \end{aligned}$$

$\mathcal{R} \triangleq (R_1, \dots, R_r)$ and F_j as defined in (23). Applying extended Schur complements on (25b) makes problem (24) and (25) equivalent.

The optimization problem (25) is a nonconvex optimization problem because of the terms $R_j = P_j^{-1}$ in (25c). We address this problem with the cone complementarity linearization algorithm ([3]) over a bisection algorithm. The algorithm is omitted for brevity; an example can be found in [7].

IV. DESIGN TRADE-OFFS

In this work, we explore the trade-off between estimation performance versus jump estimator complexity. Since the gains are related to θ_t , the solution of the previous section leads to a number of non zero different gain matrices equal to ⁵ $|\mathfrak{L}| = (\bar{d} + 1)!^{n_y} ((\bar{d} + 2)^{n_y} - 1)$ being \mathfrak{L} the non zero gain matrices fulfilling $\mathcal{L} = \mathfrak{L} \cup \{0\}$ (see (20)). We can reduce the

⁵ $|\mathfrak{L}|$ denotes the cardinal of the set \mathfrak{L} , i.e., the number of elements of \mathfrak{L} .

observer complexity by imposing some equality constraints over the set \mathcal{L} as $L_i = L_j$ in problem (25). Reducing the number of gains simplifies the numerical burden of (25), as the number of decision variables are shortened. To implement an observer with a simple online look-up-table procedure and low storage requirements, we propose the following preconfigured sets of equalities over the possible historical measurement transmission outcomes Θ (see (10)):

- **S1.** The observer gain is independent of the measurement scenario (cf. [11]), $|\mathcal{L}_{S1}| = 1$.
- **S2.** The observer gains depend on the number of real sensors from which measurements arrive successfully at each instant, $|\mathcal{L}_{S2}| = n_y$.
- **S3.** The observer gains depend on the number of real and fictitious sensors from which measurements arrive successfully at each instant, $|\mathcal{L}_{S3}| = \bar{n}_y$.
- **S4.** The observer gains depend on the measurement reception at a given instant α_t (see (13)), $|\mathcal{L}_{S4}| = 2^{\bar{n}_y} - 1$.
- **S5.** The observer gains are related to the historical measurement transmission outcomes θ_t , $|\mathcal{L}_{S5}| = (\bar{d} + 1)!^{n_y} ((\bar{d} + 2)^{n_y} - 1)$.

These gain grouping approaches, allow us to trade-off between implementation cost and estimation performance. S1 leads to the lowest cost and largest estimation error covariance, S5 gives the highest cost and best performance. The example section explores this idea.

Remark 1: [4] proposed a gain that jump with the possible instant and arrival delay for each measurement in a finite set. Adapting their proposal to ours and considering Example 1, would lead to $\mathcal{L} = \{[0 \ 0], [l_1 \ 0], [0 \ 0], [l_1 \ 0], [0 \ l_2], [l_1 \ l_2]\}$, with $l_1, l_2 \in \mathbb{R}^{2 \times 1}$ decision variables. Defining $L_1 = [l_1 \ l_2]$ and extending to the multisensor case, the method is equal to case S2.

Example 2: Considering Example 1, the proposed scenarios will impose $\mathcal{L}_{S1} = \mathcal{L}_{S2} = \{0, L_1, 0, L_1, L_1, L_1\}$, $\mathcal{L}_{S3} = \{0, L_1, 0, L_1, L_1, L_5\}$, $\mathcal{L}_{S4} = \{0, L_1, 0, L_1, L_4, L_5\}$, $\mathcal{L}_{S5} = \{0, L_1, 0, L_3, L_4, L_5\}$.

V. EXAMPLES

We consider the following system (randomly chosen)

$$A = \begin{bmatrix} 0.73 & -0.42 \\ 0.42 & 0.73 \end{bmatrix} + \rho, \quad B_w = \begin{bmatrix} 0.01 & 0.13 \\ 0.01 & 0.08 \end{bmatrix},$$

$$C = \begin{bmatrix} 0.53 & 0.39 \\ 0.72 & 0.35 \end{bmatrix},$$

with $B_u = [-0.33 \ 0.34]^T$, and where $0 \leq \rho \leq 0.5$. ρ makes the maximum absolute eigenvalue of A (denoted by $|\lambda(A)|_{\max}$) vary between $0.8422 \leq |\lambda(A)|_{\max} \leq 1.5013$. The state disturbance and sensor noises covariances are

$$W = \begin{bmatrix} 0.26 & -0.003 \\ -0.003 & 0.25 \end{bmatrix}, \quad \begin{bmatrix} \sigma_1^2 \\ \sigma_2^2 \end{bmatrix} = \begin{bmatrix} 0.0086 \\ 0.0079 \end{bmatrix}.$$

The measurements are independently acquired through a communication network that induces a delay that varies between 0 and 1. Thus, the amount of fictitious sensors is 4, $|\Theta| = (1 + 2)!^2 = 36$ (see (10)), and $|\Xi| = 2^4 = 16$ (see (14)). The probabilities of delivering a measurement with a given delay

are $\beta_1 = [0.32 \ 0.22 \ .46]$ and $\beta_2 = [0.22 \ 0.32 \ .46]$ (where $\beta_s = [\beta_{s,0} \ \dots \ \beta_{s,\bar{d}} \ \bar{\beta}_{s,\bar{d}}]$, with $s = 1, 2$).

Let us compare the results of the implementation of the optimal Kalman filter algorithm for model (5)-(6) (adapted from [11]) and the proposed algorithm. Let us define $P = C_x \mathbf{E}\{\tilde{x}_t \tilde{x}_t^T\} C_x^T$, where $C_x = [I_n \ 0_{n \times (n \cdot \bar{d})}]$ selects the covariance corresponding to $x[t] - \hat{x}[t|t]$. Then, let us introduce

$$\varepsilon(\%) = \frac{\text{tr}(P_{\text{Kal}} - P_S)}{\text{tr}(P_{\text{Kal}})} \cdot 100$$

as the factor that indicates how large the performance loss is for a given strategy S (P_S) w.r.t the one obtained with the optimal Kalman filter (P_{Kal}).

Fig. 3 and Table I show that performance gets worse when $|\lambda(A)|_{\max}$ increases its value. For a stable open-loop system, a good trade-off between performance and storage requirement can be to choose case S1, where a single gain leads to an estimation performance no more than 15% worse than the optimum. However when the system is unstable, a reasonable trade off could be to choose case S3, where with 4 gains the performance is at most 19% worse than the optimum. In the present case, the Kalman filter needs at most 976 floating-point operations per instant (including matrix inversion), while the off-line methods only need 64, which implies a reduction of a 93% in the online computing cost.

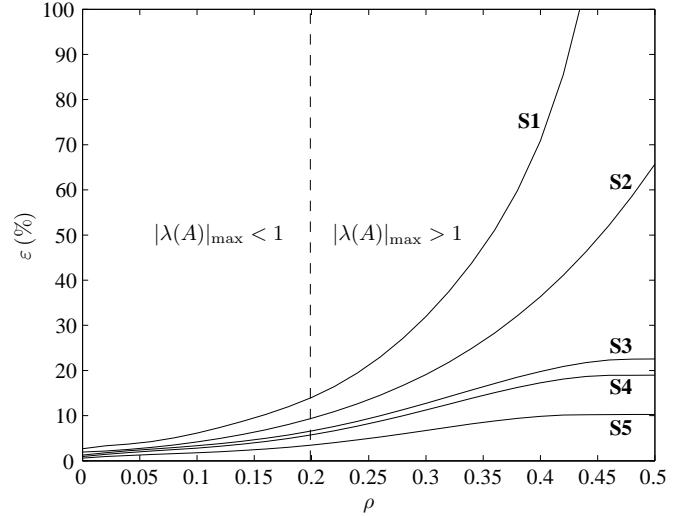


Fig. 3. Degradation of the estimation performance w.r.t the optimal Kalman filter.

TABLE I
OBSERVERS COMPARISON.

Case	S1	S2	S3	S4	S5
$ \mathcal{L} $	1	2	4	15	32
ε for $\rho = 0$	2.7	1.9	1.3	0.9	0.05
ε for $\rho = 0.5$	231.6	65.7	22.5	18.9	10.3

VI. CONCLUSIONS

In this work we develop a model for multisensor networked estimation with time-varying delays and dropouts. We introduce a Markovian finite process that stores the measurement

transmission outcomes on an interval, capturing the behavior of the network. Using this process, we design a jump state estimator for networked systems where its complexity can be chosen as a trade-off between estimation performance and storage requirements. The result is a finite set of gains that can be constrained to be equal for different values of the finite measurement outcomes parameter. Numerical results confirm that the computational cost of the on-line implementation can be much lower than Kalman filter approaches, while the achieved estimation performance is close to the optimum.

Further research may include studying Markovian delays, determining a priori the feasibility of problem (25) and analytical characterization of the performance and complexity trade-offs.

ACKNOWLEDGEMENTS

This work has been funded by MICINN project number DPI2011-27845-C02-02, and grants PREDOC/2011/37 and E-2013-02 from *Universitat Jaume I*

REFERENCES

- [1] P. Brémaud. *Markov Chains*. Springer, New York, 1999.
- [2] J. Chen, K. H. Johansson, S. Olariu, I. Ch. Paschalidis, and I. Stojmenovic. Guest editorial special issue on wireless sensor and actuator networks. *IEEE Trans. Autom. Control*, 56(10):2244–2246, 2011.
- [3] L. El Ghaoui, F. Oustry, and M. AitRami. A cone complementarity linearization algorithm for static output-feedback and related problems. *IEEE Trans. Autom. Control*, 42(8):1171–1176, 1997.
- [4] C. Han, H. Zhang, and M. Fu. Optimal filtering for networked systems with markovian communication delays. *Automatica*, 49(10):3097 – 3104, 2013.
- [5] J. P. Hespanha, P. Naghshtabrizi, and Y. Xu. A survey of recent results in networked control systems. *Proceedings of the IEEE*, 95(1):138–162, 2007.
- [6] X. Liu and A. Goldsmith. Kalman filtering with partial observation losses. In *Proc. Conf. Decision. and Control*, pages 4180–4186, 2004.
- [7] I. Peñarocha, D. Dolz, and R. Sanchis. Inferential networked control with accessibility constraints in both the sensor and actuator channels. *International Journal of Systems Science*, 2013.
- [8] I. Peñarocha, R. Sanchis, and P. Albertos. Estimation in multisensor networked systems with scarce measurements and time varying delays. *Systems & Control Letters*, 61(4):555–562, 2012.
- [9] I. Peñarocha, R. Sanchis, and J. A. Romero. State estimator for multisensor systems with irregular sampling and time-varying delays. *International Journal of Systems Science*, 43(8):1441–1453, 2012.
- [10] M. Sahebsara, T. Chen, and S. L. Shah. Optimal \mathcal{H}_2 filtering in networked control systems with multiple packet dropout. *IEEE Trans. Autom. Control*, 52(8):1508–1513, 2007.
- [11] L. Schenato. Optimal estimation in networked control systems subject to random delay and packet drop. *IEEE Trans. Autom. Control*, 53(5):1311–1317, 2008.
- [12] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry. Kalman filtering with intermittent observations. *IEEE Trans. Autom. Control*, 49(9):1453–1464, 2004.
- [13] S. C. Smith and P. Seiler. Estimation with lossy measurements : jump estimators for jump systems. *IEEE Trans. Autom. Control*, 48(12):2163–2171, 2003.

APPENDIX

Equation (21) is obtained using the law of total probabilities. Considering the independency between x_{t-1} , \bar{v}_t and w_{t-1} ,

$P_{t,j} = \mathbf{E}\{\tilde{x}_t \tilde{x}_t^T | \theta_t = \vartheta_j\}$ can be calculated as follows.

$$\begin{aligned} & \sum_{i=0}^r \Pr\{\theta_{t-1} = \vartheta_i | \theta_t = \vartheta_j\} \mathbf{E}\{\tilde{x}_t \tilde{x}_t^T | \theta_{t-1} = \vartheta_i, \theta_t = \vartheta_j\} = \\ & = \sum_{i=0}^r p_{i,j} \frac{\pi_i}{\pi_j} F_j (\bar{A} \mathbf{E}\{\tilde{x}_{t-1} \tilde{x}_{t-1}^T | \theta_{t-1} = \vartheta_i\} \bar{A}^T + \mathbf{E}\{w_{t-1} w_{t-1}^T\}) F_j^T \\ & + \sum_{i=0}^r p_{i,j} \frac{\pi_i}{\pi_j} X_j \mathbf{E}\{\bar{v}_t \bar{v}_t^T\} X_j^T \end{aligned}$$

which leads to (22) after using $\Pr\{\theta_{t-1} = \vartheta_i | \theta_t = \vartheta_j\} = \Pr\{\theta_t = \vartheta_j | \theta_{t-1} = \vartheta_i\} \Pr\{\theta_{t-1} = \vartheta_i\} / \Pr\{\theta_t = \vartheta_j\}$.

Considering the linear operator in Lemma 1, Theorem 1 and constraint (24b), we have $\mathcal{T}(\bar{\mathcal{P}}) \prec \mathfrak{E}\{\bar{\mathcal{P}}\} \preceq \bar{\mathcal{P}}$. Thus, $\mathcal{T}(\cdot)$ meets the condition of Lemma 1. The evolution of \mathcal{P}_t is expressed as $\mathcal{P}_{t+1} = \mathfrak{E}\{\mathcal{P}_t\} = \mathcal{T}(\mathcal{P}_t) + U$. Since U contains the disturbance and noise covariance (both positive definite and bounded), then $U \succ 0$, leading that $\{\mathcal{P}_t\}$ is bounded.

First, let us show the convergence of sequence $\{\mathcal{P}_t\}$ with initial value $\mathcal{Q}_0 = 0$, where $\mathcal{Q}_t \triangleq (\mathcal{Q}_{t,0}, \dots, \mathcal{Q}_{t,r})$. Let $\mathcal{Q}_t = \mathfrak{E}\{\mathcal{Q}_{t-1}\} = \mathfrak{E}^t\{\mathcal{Q}_0\}$, then from (22), $\mathcal{Q}_1 \succeq \mathcal{Q}_0 = 0$ and $\mathcal{Q}_1 = \mathfrak{E}\{\mathcal{Q}_0\} \preceq \mathfrak{E}\{\mathcal{Q}_1\} = \mathcal{Q}_2$. By induction, $\{\mathcal{Q}_t\}$ is non decreasing. Also, by Lemma 1, $\{\mathcal{Q}_t\}$ is bounded and by Theorem 2 there exists an $M_{\mathcal{Q}} \triangleq (M_{\mathcal{Q},0}, \dots, M_{\mathcal{Q},r})$ such that $\mathcal{Q}_t \preceq M_{\mathcal{Q}}$ for any t . Hence, the sequence converges and $\lim_{k \rightarrow \infty} \mathcal{Q}_t = \bar{\mathcal{P}} \succeq 0$, where $\bar{\mathcal{P}}$ is a fixed point, i.e. $\bar{\mathcal{P}} = \mathfrak{E}\{\bar{\mathcal{P}}\}$. Second, we state the convergence of $\mathcal{G}_t = \mathfrak{E}^k\{\mathcal{G}_0\}$, initialized at $\mathcal{G}_0 \succeq \bar{\mathcal{P}}$ where $\mathcal{G}_t \triangleq (G_{t,0}, \dots, G_{t,r})$. Since $\mathcal{G}_1 = \mathfrak{E}\{\mathcal{G}_0\} \succeq \mathfrak{E}\{\bar{\mathcal{P}}\} = \bar{\mathcal{P}}$, then $\mathcal{G}_t \succeq \bar{\mathcal{P}}$ for any t . Moreover $0 \preceq \mathcal{G}_{t+1} - \bar{\mathcal{P}} = \mathfrak{E}\{\mathcal{G}_t\} - \mathfrak{E}\{\bar{\mathcal{P}}\} = \mathcal{T}(\mathcal{G}_t - \bar{\mathcal{P}})$. As $\mathcal{G}_t - \bar{\mathcal{P}} \succeq 0$, following the results on Lemma 1, then $0 \preceq \lim_{t \rightarrow \infty} (\mathcal{G}_t - \bar{\mathcal{P}}) = 0$, i.e., the sequence $\{\mathcal{G}_t\}$ converges to $\bar{\mathcal{P}}$.

We demonstrate now that for any initial condition $\mathcal{P}_0 \succeq 0$, the iteration $\mathcal{P}_t = \mathfrak{E}\{\mathcal{P}_{t-1}\}$ converges to $\bar{\mathcal{P}}$. Since $0 \preceq \mathcal{Q}_0 \preceq \mathcal{P}_0 \preceq \mathcal{G}_0$, we derive by induction that $0 \preceq \mathcal{Q}_t \preceq \mathcal{P}_t \preceq \mathcal{G}_t$. Therefore, as $\{\mathcal{Q}_t\}$ and $\{\mathcal{G}_t\}$ converge to $\bar{\mathcal{P}}$, then $\{\mathcal{P}_t\}$ also converges to $\bar{\mathcal{P}}$ and the convergence is demonstrated. Finally, we need to show that

$$\bar{\mathcal{P}} = \arg \min_{\bar{\mathcal{P}}} \text{tr} \left(\sum_{j=0}^r P_j \pi_j \right) \text{ subject to (24b).}$$

Suppose this is not true, i.e. $\hat{\mathcal{P}}$ solves the optimization problem, but $\hat{\mathcal{P}} \neq \mathfrak{E}\{\hat{\mathcal{P}}\}$. Since $\hat{\mathcal{P}}$ is a feasible solution, then $\hat{\mathcal{P}} \succ \mathfrak{E}\{\hat{\mathcal{P}}\} = \hat{\hat{\mathcal{P}}}$. However, this implies $\text{tr} \left(\sum_{j=0}^r \hat{P}_j \pi_j \right) > \text{tr} \left(\sum_{j=0}^r \hat{\hat{P}}_j \pi_j \right)$, which contradicts the hypothesis of optimality of matrix $\hat{\mathcal{P}}$. Therefore $\hat{\mathcal{P}} = \mathfrak{E}\{\hat{\mathcal{P}}\}$. Furthermore $\bar{\mathcal{P}}$ is unique since for a set of observer gains such that

$$[\bar{\mathcal{P}}, \mathcal{L}] = \arg \min_{\bar{\mathcal{P}}, \mathcal{L}} \text{tr} \left(\sum_{j=0}^r P_j \pi_j \right) \text{ subject to (24b),}$$

we have shown that the sequence converges to $\bar{\mathcal{P}}$, and this concludes the theorem.